DELPHI PROCESS: A Methodology Used for the Elicitation of Opinions of Experts

Bernice B. Brown*

The RAND Corporation, Santa Monica, California

---

*Any views expressed in this paper are those of the author. They should not be interpreted as reflecting the views of The RAND Corporation or the official opinion or policy of any of its governmental or private research sponsors. Papers are reproduced by The RAND Corporation as a courtesy to members of its staff.

This paper was prepared for publication in ASTME Vectors, February 1968.

DELPHI PROCESS: A Methodology Used for the Elicitation of Opinions of Experts

Bernice B. Brown

The RAND Corporation, Santa Monica, California

Introduction

Forecasting is a fundamental part of planning in the modern world. A forecaster is one who, after observing a series of historical facts, undertakes to guess some related future event. His materials are the available data about the past; facts about some specified phenomenon in several past epochs. Using the traditional statistical techniques of forecasting, he considers the facts and arrives at some opinion as to the future course of this phenomenon. The forecaster is not a prophet but he is a believer in trends and he has some faith in the continuity of nature.

Short-term forecasts of events that may occur next week or next month are usually accurate enough to be useful. But the uncertainties multiply rapidly when long-range forecasts are required and many of the planning decisions must of necessity be based on a series of intuitive judgments.

Intuition and judgment permeate all analysis, not only as to which hypotheses should be tested or what facts are relevant but also in supplementing a model of a process when the quantitative mathematical model is known to be inadequate. It is inevitable that as questions to be answered get broader and more complex, intuition and judgment must supplement quantitative analysis to an increasing extent. The Delphi process

is representative of an important class of techniques ~~that~~ need to be developed for further applications to decision making situations. It involves one of the methodological aspects of modern practice in operations research, namely the reliance on judgment of experts.

For many years experts have been used in brainstorming sessions and round-table discussion groups with the object of achieving a group opinion, a group solution to a problem or a group estimate of some unknown numerical quantity. The traditional discussion approach was often beset by psychological factors such as the presence of a dominant, persuasive personality, the tendency to want to meet the approval of the group and the unwillingness to change an opinion which had been publicly expressed.

Helmer and Rescher [1] point out that the Delphi Technique "eliminates committee activity altogether, thus further reducing the influence of certain psychological factors, such as specious persuasion, the unwillingness to abandon publicly expressed opinions, and the bandwagon effect of majority opinion. This technique replaces direct debate by a carefully designed program of sequential individual interrogations (best conducted by questionnaires) interspersed with information and opinion feedback derived by computed consensus from the earlier parts of the program. Some of the questions directed to the respondents may, for instance, inquire into the reasons for previously expressed opinions and a collection

of such reasons may then be presented to each respondent in the group, together with an invitation to reconsider and possibly revise his earlier estimates. Both the inquiry into the reasons and subsequent feedback of the reasons adduced by others may serve to stimulate the experts into taking into due account considerations they might through inadvertence have neglected, and to give due weight to factors they were inclined to dismiss as unimportant on first thought."

Description of Delphi

The Delphi method is a name that has been applied to a technique used for the elicitation of opinions with the object of obtaining a group response of a panel of experts. Delphi replaces direct confrontation and debate by a carefully planned, orderly program of sequential individual interrogations usually conducted by questionnaires. The series of questionnaires is interspersed with feedback derived from the respondents. Respondents are also asked to give reasons for their expressed opinions and these reasons are subjected to a critique by fellow respondents. The technique puts the emphasis on informed judgment. It attempts to improve the panel or committee approach by subjecting the views of individual experts to each other's criticism in ways that avoid face to face confrontation and provide anonymity of opinion and of arguments advanced in defense of those opinions.

The first step in the application of the Delphi method is the selection of a group of experts. Wise decision makers

have always depended upon the advice of experts but often the consultation with specialists has been haphazard and there has been no attempt to collate differences of opinion among the experts. The selection of experts is an intricate problem even when the category of expertise needed is well-defined. A man's expertness might be judged by his status among his peers, by his years of professional experience, by his own self-appraisal of relative competence in different areas of inquiry, by the amount of relevant information to which he has access or by some combination of objective indices and a priori judgment factors.

Suppose the question to be answered is one involving a forecast of a numerical quantity, namely "What will the world population be in the year 2000?" The panel of experts in such a case might consist of specialists in sociology, demography and population growth. This question was imbedded in one part of the panel on Population Control in Helmer's "Long-Range Forecasting Study" [2]. I have simplified it for use as an illustrative example. The procedure will be described by pursuing the response to this question through the successive questionnaires.

In the first questionnaire, all respondents would be asked to record their estimate of the world population in 2000. Each respondent would also be asked to assign a number 1, 2, 3, or 4 as a relative rating, using 1 for the relatively most competent. This score would constitute a self-appraisal.

A respondent would be expected to look at all of the questions in the set and assess his relative competency on each one. The information from these responses which would furnish feedback data for the second interrogation would be the median and the interquartile range (i.e., the middle 50 percent of the responses).

In the second round, respondents would be asked to reconsider their estimate and revise it if they desired. They would also be asked to give the reasons for the estimate and state what factors were considered in obtaining the answer. They may also be asked to describe the rationale that led them to a revision of their original estimate. Some of the reasons given for population estimates at the low end of the scale were (a) rapid increase in use and effectiveness of birth control measures, (b) increased economic prosperity, (c) progress in welfare and education in the developing nations and (d) attrition due to war and disease Among the reasons for high estimates were (a) medical advances resulting in lower death rates, (b) insufficient acceptance of birth control measures, (c) development of centralized world government providing efficient distribution of food, shelter and services and (d) advances in agriculture. Participants indicated that they projected birth rates and death rates and net growth rates in arriving at the population estimates.

In the third questionnaire, the median and interquartile range of the previous round would be given along with a summary of reasons for high and low population estimates.

Participants would be asked to give a critique of the reasons offered by members of the group and to specify which arguments were found to be unconvincing and why. Responses to the third round included estimates that the death rate would drop from 19 per 1000 to a figure between 10 and 17 per 1000 and that birth rates would decline from 36 per 1000 to a figure between 15 and 26 per 1000.

In the fourth round the median and interquartile range of the previous round would again be used as numerical feedback. The counter-arguments against reasons for high and low estimates would be summarized. Majority and minority opinions on the projection of death rates and birth rates would be described and respondents then asked to reconsider the pros and cons presented and give a final, possibly revised, estimate of the world population in the year 2000. Each respondent would also be given an opportunity to revise his own relative competence rating.

The median of these final responses would then be taken to represent the group response on the required answer.

Some Modifications of the Procedure

The objective of a given inquiry and the special problems associated with the area of expertise being tapped might suggest a number of modifications or refinements. In the illustrative example of obtaining an estimate of world population in 2000, the respondents could be asked to suggest subsidiary questions whose answers would be helpful in formulating the estimate.

Respondents might ask some of the following questions: What is the world population at the present time? What was the rate of increase of population during the last 50 years? What is the expected length of life at age 1 at the present time in the U.S.? What percent of the 1900 world population were 18 years old or less? What percent were 65 or older?

Answers to these subsidiary questions could be solicited from the group and fed back to the participants. It would also be possible to use a member of the control team administering the experiment or an outside specialist as a "resource analyst," and answers to the subsidiary questions could be researched and passed to the participants as supplementary information along with the citation of the reference source. This procedure could become cumbersome if the number of questions under consideration were sizable.

In the course of a few years, it should be possible to equip each expert with a console through which he could feed his responses into a computer. The computer would process them, compute some measures of the group response, possibly add relevant information from an existing data bank and feed the results back to each respondent. At the Rand Corporation, small Delphi experiments are being conducted that use a number of personal electric typewriter consoles connected through an on-line time-sharing computer system.

How Delphi Method has been Used

A small experiment called "Twenty Questions" [3] was conducted at Rand in 1964 using staff members as participants. The

Participants would be asked to give a critique of the reasons offered by members of the group and to specify which arguments were found to be unconvincing and why. Responses to the third round included estimates that the death rate would drop from 19 per 1000 to a figure between 10 and 17 per 1000 and that birth rates would decline from 36 per 1000 to a figure between 15 and 26 per 1000.

In the fourth round the median and interquartile range of the previous round would again be used as numerical feedback. The counter-arguments against reasons for high and low estimates would be summarized. Majority and minority opinions on the projection of death rates and birth rates would be described and respondents then asked to reconsider the pros and cons presented and give a final, possibly revised, estimate of the world population in the year 2000. Each respondent would also be given an opportunity to revise his own relative competence rating.

The median of these final responses would then be taken to represent the group response on the required answer.

Some Modifications of the Procedure

The objective of a given inquiry and the special problems associated with the area of expertise being tapped might suggest a number of modifications or refinements. In the illustrative example of obtaining an estimate of world population in 2000, the respondents could be asked to suggest subsidiary questions whose answers would be helpful in formulating the estimate.

Respondents might ask some of the following questions: What is the world population at the present time? What was the rate of increase of population during the last 50 years? What is the expected length of life at age 1 at the present time in the U.S.? What percent of the 1900 world population were 18 years old or less? What percent were 65 or older?

Answers to these subsidiary questions could be solicited from the group and fed back to the participants. It would also be possible to use a member of the control team administering the experiment or an outside specialist as a "resource analyst," and answers to the subsidiary questions could be researched and passed to the participants as supplementary information along with the citation of the reference source. This procedure could become cumbersome if the number of questions under consideration were sizable.

In the course of a few years, it should be possible to equip each expert with a console through which he could feed his responses into a computer. The computer would process them, compute some measures of the group response, possibly add relevant information from an existing data bank and feed the results back to each respondent. At the Rand Corporation, small Delphi experiments are being conducted that use a number of personal electric typewriter consoles connected through an on-line time-sharing computer system.

How Delphi Method has been Used

A small experiment called "Twenty Questions" [3] was conducted at Rand in 1964 using staff members as participants. The

purpose was to get some insight into the methodology of the Delphi process, and the questions submitted to participants were World Almanac type questions for which numerical answers were available. The experiment was designed to determine to what extent initially divergent opinions of a group of persons would converge if successive questioning were re-enforced by a feedback process in which respondents were given information from other members of the group. The set of respondents was divided into two groups, a primary group and a control group. The primary group completed four questionnaires with feedback information. The control group participated in two rounds only, with no feedback at all. The results of the experiment indicated that a sharper consensus was obtained by the primary group, that the total range of the responses was reduced on successive rounds but the accuracy of the group estimate was no greater with the primary group than with the control if we compare them on the response to the second round. It is necessary to state what criterion was used to measure accuracy. In this case, it was the sum of the absolute values of the logarithm of x/T, where x is the group median response and T is the true answer. It was evident that a reconsideration of the questions fostered convergence in both the primary and the control groups.

An experiment in which the Delphi technique was applied was reported by Dalkey and Helmer [4]. The results of the experiment were released for open publication in 1962 but the experiment itself was conducted about ten years earlier. In this application, expert opinion was applied to the dual

problem of the selection of an optimal industrial target system and the estimation of the number of A-bombs required to reduce the munitions output by a prescribed amount. Seven experts participated, responding to five questionnaires submitted at approximately weekly intervals. The first questionnaire was followed by an interview in which each respondent was asked to reproduce the reasoning by which he arrived at an estimate of the number of bombs and to show the component breakdown by industries. The third also was followed by an interview for the clarification of ambiguities. The choices of target systems were quite distinct, the only common feature being the inclusion of the steel industry in each. The numerical quantity being estimated showed considerable convergence. The ratio between the largest and smallest response was about 100 to 1 on the initial round but had dropped to about 3 to 1 on the final round.

A first application of the Delphi technique to long-range forecasting was made in 1964 [2] in an investigation in which several panels of experts were asked to make contingency forecasts of the state of the world twenty-five to fifty years hence. Six groups of experts were selected, one for each area of inquiry. Each panel answered four sequential questionnaires spaced approximately two months apart. The six areas covered were scientific breakthroughs, population growth, automation, space progress, war prevention and future weapon systems. The six groups of respondents made judgments on more than 200

predictive items during the course of the experiment. The substantive outcome of the investigation cannot be summarized here. A few thought-provoking examples of some of the predictions that were articulated by the experts have been selected from pages 45-46 of the Gordon and Helmer report [2]. (1) The implication that the water-covered portions of the earth may become important enough to warrant national territorial claims; (2) the possibility that continued developments in automation will result in serious social upheavals and the almost complete acceptance of the necessity of regulative legislation; (3) the strong likelihood of the emergence of weapons of a nonkilling, nonproperty-destroying nature, covert perhaps, attacking on the psychological or biological level; (4) the eventual abundance of resources of energy, food, and raw materials, but also the possibility that a continuing inequitable world distribution of these assets to the increasing world population may furnish a persisting stimulant to warfare.

A study, entitled Innovation in Education, was carried out at the Institute of Government and Public Affairs at UCLA under the sponsorship of the Charles F. Kettering Foundation during 1966. The report was published by authors Adelson, Alkin, Carey and Helmer [5]. This study was an attempt by a multi-disciplinary group to generate some perspectives on possible changes in american education. The Delphi technique was included as part of the investigations because the researchers conjectured that it might be useful as a planning aid for those who make educational policy. Respondents examined and made

judgments on a list of nearly one hundred proposed educational innovations covering a wide range of educational activities. Proposed innovations covered Curriculum, Teaching Methods, School Administration, Student Participation, Staff Utilization, Adult Retraining, Automated Education, etc. In the preparation of the final round, each of the proposed items was assigned to a gross cost category. If a respondent thought the cost category assignment for a given innovation was inappropriate, he was expected to state the reason for his position. However, the principal task for the respondents in the final questionnaire was to allocate a fictitious 5 year budget of ten billion dollars among the proposed innovations. The authors state that the participants found it difficult to make the required choices even though they were well informed in the field and were used to making decisions. The results of the study indicated that the Delphi technique may be potentially useful in educational planning.

Robert M. Campbell[6] used the Delphi technique in a study in which business and economic indices were forecast. He conducted a controlled experiment using students in two graduate seminars in business forecasting. Each seminar was divided at random into two equal groups. All four groups were asked to make forecasts of 16 economic series for the first quarter of 1966. The participants were informed of the experiment about a month in advance and were given some guidance on accumulating information which would develop their expertise without revealing the actual series to be forecast. One group

in each seminar used the traditional methods of making business forecasts and the other used the Delphi process. The traditional method allowed participants to interact freely with others in the group for the purpose of obtaining information relative to the forecasts. The Delphi experimental group gave individual responses to a series of four questionnaires over a period of six weeks. The group participants who used the Delphi process made more accurate forecasts than the group using the traditional business forecasting technique.

A study was conducted within TRW, Inc. [7] in an attempt to predict the operating environment of the company twenty years hence. The method used was to ask each member of a panel of 27 technologists to list events of a technical nature that were likely to occur within the next 20 years. Participants were from all working groups in the company and each man was expected to suggest events that might have substantial impact on potential product lines of his group. The lists of technological breakthroughs were collected by mail. These were compiled and the completed document was returned to each panelist with the suggestion that he should edit freely in his own area of expertise. The TRW probe of the future resulted in a list of about 400 events with predicted dates of occurrence representing the judgment of responsible experts in several areas of research. The results constitute an information source for planners throughout the corporation.

Potential Applications

The judgment of experts may be called on in any planning operation in which it is necessary to choose among several alternative courses of action and no theory has been developed which would evaluate the consequences of the proposed courses of action with one course singled out as the preferred alternative by traditional maximization procedures. We use an expert because he has at his disposal a large store of background knowledge and a cultivated sensitivity to its relevance which permeates his intuitive insight. We need a consensus of experts because individual experts will disagree and we are unwilling to rely on the judgment of a single specialist.

There are many examples of the use of expert judgment for prediction. One example is provided by the field of medical diagnostics. Another is in the use of the advice of an expert investment counselor.

There are some indications that the Delphi process would be useful as a business forecasting tool. Market forecasts are often judgment forecasts and the group response arrived at by the Delphi procedure might prove to be more accurate than reliance on any one individual. Several areas of industrial forecasting ranging from financial planning to sales prediction may be fertile areas. Some of the management decisions made in the promotion and distribution of products and in product pricing problems might be enriched by information obtained from the Delphi procedure.

Often there are variables used as inputs to models for which no adequate measure exists. An example might be a policy decision model in which measures of social and cultural conditions in some of the developing nations are needed. There are no historical records available. The obvious recourse is the efficient use of the intuition and judgment of a group of persons who are keen observers and have lived in the country for a long period of time.

The use of expertise is not a retreat from objectivity. Judgment and informed opinion have always played a crucial role in human enterprises. Expert judgment can be incorporated into the structure of an investigation and can be made subject to some of the safeguards that are commonly used to assure objectivity in any scientific inquiry.

REFERENCES

1. Helmer, O., N. H. Rescher, On the Epistemology of the Inexact Sciences, The RAND Corporation R-353, February 1960. Also, Management Science, Vol. 6, 1959, pp. 25-52.

2. Gordon, T., O. Helmer, Report on a Long-range Forecasting Study, The RAND Corporation P-2982, September 1964. Also, O. Helmer, T. Gordon, and B. Brown, Social Technology, Basic Books, New York, 1966.

3. Brown, B., O. Helmer, Improving the Reliability of Estimates Obtained from a Consensus of Experts, The RAND Corporation P-2986, September 1964. Also, O. Helmer, Social Technology, Basic Books, New York, 1966.

4. Dalkey, N., O. Helmer, An Experimental Application of the Delphi Method to the Use of Experts, The RAND Corporation RM-727, July 1962. Also, Management Science, Vol. 9, 1963, pp. 458-467.

5. Adelson, M., M. C. Alkin, C. Carey, and O. Helmer, "Planning Education for the Future: Comments on a Pilot Study," American Behavioral Scientist, March 1967, pp. 1-12, 21-31.

6. Campbell, Robert M., A Methodological Study of Expert Utilization in Business Forecasting. Unpublished doctoral dissertation, University of California, Los Angeles, 1966.

7. North, H. Q., TRW Looks at the Future, TRW Systems, Inc., 1967.